基于情感分析的网络谣言识别方法*

首欢容　邓淑卿　徐　健
(中山大学资讯管理学院　广州 510006)

摘要: 【目的】提出一种基于情感分析技术自动识别特定领域谣言的方法。【方法】界定高、低质量信息源, 在假设高质量信息源信息更可靠的情况下, 通过基于情感词典的情感分析方法, 量化高质量信息源与低质量信息源对特定对象的情感差异, 判定低质量信息源提供的信息是否属于谣言。【结果】将该方法应用于"食品养生"、"医学健康"两个领域进行谣言识别。在 30 个疑似谣言案例中准确识别出 23 个谣言案例, 准确率为 76.67%。本文提出的谣言识别方法在谣言预测方面的 F 值为 83.34%, 查全率为 71.42%, 查准率为 100%; 在非谣言文本预测上的 F 值为 72.73%, 查全率为 100%, 查准率为 57.14%。【局限】未实现不同信息源数据自动抽取, 每个谣言案例下的人工收集的谣言数量有限。【结论】本文基于情感分析的谣言识别方法对特定类型的谣言是有效的。

关键词: 情感分析　情感词典　谣言检测　谣言识别
分类号: G350

1 引　言

谣言是一种普遍的社会舆论现象[1], 不同学者对谣言的具体内涵给予不同定义:《现代汉语词典》中对谣言的定义为"没有事实根据的消息"[2]。沙莲香[3]将谣言定义为: "一种来路不明的、传无根据的、内容没有得到确认的, 缺乏事实根据的信息"。总体而言, 缺乏事实依据的谣言具有强烈说服他人相信某种信息的目的, 往往采用夸张语言风格, 有强烈的情绪化特征。

新的传播媒介的诞生与发展使谣言的传播速度更快、影响范围更大。谣言干扰公民已有认知, 进一步导致公众非理性行为, 危害社会安定。及时识别谣言并扼制谣言传播, 净化网络环境, 是当前网络环境下一个重要议题。如何准确、高效地识别网络谣言, 是控制谣言传播首要且关键的步骤。

情感分析技术可以识别语料的情感倾向及程度, 其在用户意见挖掘、政府民意调查等多方面都具有广泛应用[4]。由网络谣言具有夸张语言风格、异常情感特征的特点, 本文提出一种基于情感分析技术的谣言识别方法, 以不同质量信息源的文本情感冲突识别特定领域的网络谣言。

2 相关研究

2.1 情感分析方法识别技术

情感分析又称意见挖掘, 是指通过对用户发表的内容文本进行主客观观点、情绪、极性的分析和挖掘, 判断文本的情感倾向分类[5]。目前, 情感分析方法主要有基于情感词典和机器学习的方法。

(1) 基于情感词典方法的情感分析技术的核心是构建特定的情感词典, Tong[6]通过人工抽取与电影影评相关的词汇, 人工标注情感词极性, 建立专门的情感词典。中文方面, 通用词典有大连理工大学情感词汇本体库[7]、知网 HowNet[8]情感词典等; 在专用词典方面, 陈晓东[9]构建了一个面向微博的情感词典。

(2) 基于机器学习方法的情感分析技术核心是构建分类器, 对语料进行情感分类。目前机器学习的分类算法有支持向量机、朴素贝叶斯等。

在情感分析技术的应用方面, 主要应用有用户意

通讯作者: 徐健, ORCID: 0000-0003-4886-4708, E-mail: issxj@mail.sysu.edu.cn。

*本文系国家社会科学基金项目"用户评论情感分析及其在竞争情报服务中的应用研究"(项目编号: 11CTQ022)和广东省科技专项"基于内容的科技文献分析服务平台"(项目编号: 2016B030303003)的研究成果之一。

见挖掘[10]、电视节目评分预测[11]、股票市场价格预测[12]等，但是尚未见应用情感分析技术识别谣言的相关研究。

2.2　谣言识别技术

目前谣言检测主要采用机器学习方法，两个核心步骤为谣言特征提取及分类器训练，最后利用训练的分类器进行谣言识别。

毛二松等[13]提出一种基于深层特征和集成分类器的微博谣言检测方法，由于抽取的特征具有明显微博平台倾向性，该检测方法应用平台有待拓展。Qazvinian 等[14]通过提取 Twitter 谣言文本中的浅层文本特征、元素特征和行为特征，构建多个贝叶斯分类器和集成分类器，以识别 Twitter 中的谣言。Kwon 等[15]提出一种基于不同时间序列的谣言探测方法，研究结果表明在谣言传播的不同时期，选取不同特征将影响谣言识别效果。

机器学习方法的核心是构建良好分类器。分类器效果依赖于训练集及特征选择，时间与人力成本较高。本文基于情感分析技术，提出通过不同质量信息源的情感冲突来识别谣言的方法。

3　基于情感分析的谣言识别方法设计

3.1　理论依据

谣言具有一定的异常情感特征。张志安等[16]总结了谣言的主题与特征：在主题方面，谣言的叙事主题分为健康类、时政类和社会类，其中健康类主题在所有微信谣言文章数量中占主体地位；在特征方面，绝大部分谣言基本上都是针对公众感到恐惧、焦虑或担忧的议题诉诸恐怖说服，这些谣言标题内容叙述夸张、语言口语化、缺乏科学严谨性并试图采用“情感策略”，煽动读者情绪。

基于此，笔者认为谣言与非谣言之间具有情感倾向冲突，这是本文基于情感分析技术的谣言检测方法的理论依据。

3.2　方法框架设计

本文将具有以下特征的谣言定义为“简单谣言”：内容简单、缺乏复杂的逻辑推理，文本直接对某件事物做出好、坏评价，并且评价的情感倾向非常明显。本文提出基于情感分析的谣言识别技术主要用于识别简单谣言。计算不同质量信息源文本的情感值及其情感差异，如果情感差异过大，则说明高、低质量信息源对同一事物的情感倾向不同，假设高质量信息源更可信，认为低质量信息源的信息属于谣言，达到识别谣言的目的。基于情感分析技术的谣言识别方法包括 4 个模块：高低质量信息源界定、文本预处理、文本情感值计算和文本谣言识别，总体流程框架如图 1 所示。

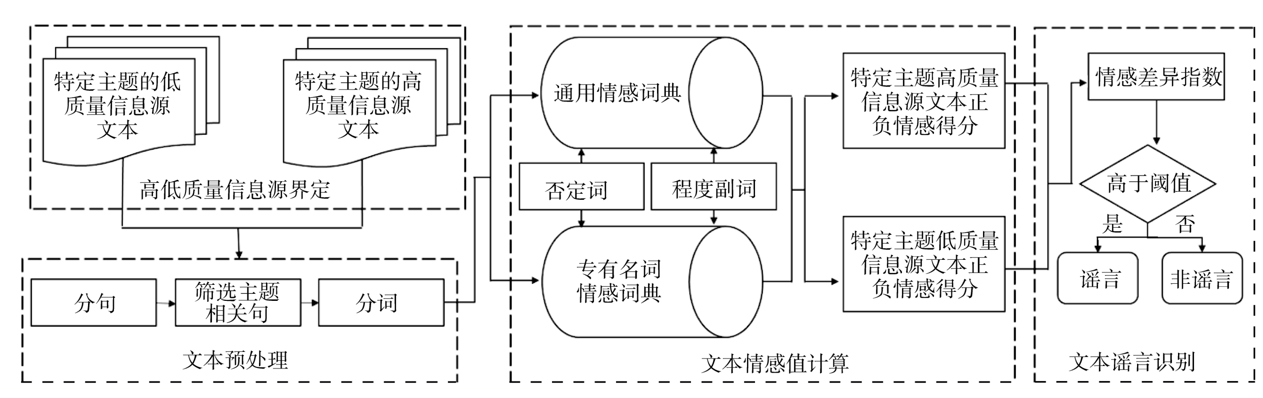


图 1　基于情感分析技术的谣言识别方法的总体框架

(1) 高、低质量信息源的界定

界定高质量信息源和低质量信息源是准确识别谣言的基础。目前，对信息源的评价一般有两种评价方法：直接评价法和间接评价法[17]。

①直接评价法一般通过建立指标评价体系的方法，对不同信息来源媒介每一项指标进行打分，综合各项指标对信息源进行评价。

②间接评价法通过信息用户来评价信息源，以调查表的方式调查用户对信息源的需求和利用情况，其评价较为客观，但是工作量大，需要信息用户的高度配合。

在实际应用中，可结合客观条件，选择合适的评价方法评价信息源，从而界定高、低质量信息源。

(2) 文本预处理

在疑似谣言文本中，其针对不同的对象可能具有

不同的情感倾向，如果计算疑似谣言文本的整体情感倾向，则会降低情感计算准确性。因此在情感计算前过滤与特定对象无关的句子，再进行分词，便于后续情感词匹配。

(3) 文本情感值计算

基于情感词典的情感分析方法关键步骤之一为构建情感词典。根据研究目的、主题的不同，研究者可以采用不同的通用情感词典及专有名词词典。专有名词是指研究不同学科或主题时会涉及的对应领域专有的名词词汇。对于专有名词，不能直接计算其情感值，需要借助情感动词词典。程度副词会影响情感程度表达，否定词会影响情感极性表达，因此在计算情感值的过程中需要识别程度副词和否定词来修正情感值，本文借用杜嘉忠等[18]提出的距离与词性的方法计算文本的情感值，具体的情感值计算流程如图 2 所示。

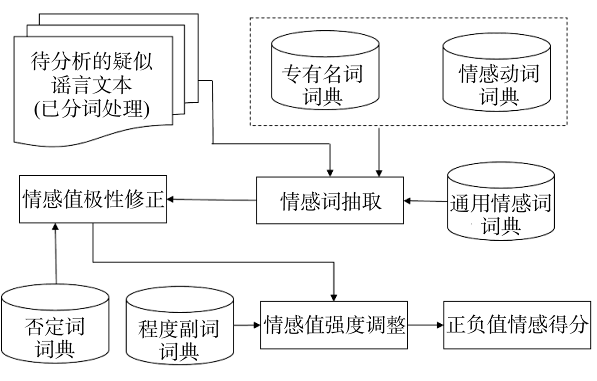


图 2 疑似谣言文本的情感值计算流程

依据图 2 流程对疑似谣言文本的情感值进行计算，最初每个谣言文本都有一个正向情感得分 *positive* 和一个负向情感得分 *negative*，其初始值均为 0；每个情感词的初始情感极性转折次数 $reverse\ Time_{SentiWord}$=0，转折标志 $reverse_{SentiWord}$=1；情感程度系数初始为 *degree*=1；对于每一个情感词 *SentiWord*，其在情感词典中的得分为 $SentiDic_{SentiWord}$；对于每一个程度词 *DegWord*，其程度系数为 $DegDic_{DegWord}$。

对于通用情感词典计算情感值，其具体流程为：遍历疑似谣言文本词表，如果词语与通用情感词典中词汇匹配，则以该情感词为中心。

①寻找并统计该情感词前 5 个词及该情感词后两个词范围内否定词的个数 $reverseTime_{SentiWord}$，即如果 $reverse\ Time_{SentiWord}$ 是奇数，修正情感极性，令 $reverse_{SentiWord}=-1$；否则不需修正情感极性。

②寻找该情感词前 5 个词语及该情感词后两个词语范围内的程度副词 *DegWord*，用 *degree* 乘以匹配程度词语在程度级别词典中的情感强度系数，如公式(1)所示。

$$degree = degree \times DegDic_{DegWord} \tag{1}$$

③分别累加对应的正、负面情感得分，规则如公式(2)所示，其中 *SentimentValue* 代表正面情感得分 *postive* 或负面情感得分 *negative*，当 $SentiDic_{SentiWord} \times reverse_{SentiWord} > 0$，情感得分累计到 *positive*，否则，情感得分累计到 *negative*。

$$SentimentValue = \sum SentiDicSentiWord \times degree \times reverseSentiWord \tag{2}$$

对于专有名词词典情感值计算，具体流程与通用情感词词典情感值计算类似。其区别在于在遍历疑似谣言文本词表时，如果词语匹配专有名词词典中的词汇，其不一定具有情感表达，需进一步查找是否具有修饰专有名词的情感动词。如果有，则情感动词的情感值为 *Score*。以专有名词词汇为中心，寻找名词的修饰词与否定词，累加最后对应的正负情感得分。具体步骤如下。

遍历疑似谣言文本词表，如果词语匹配到专有名词词典中的词汇。

①寻找该专有名词前后的词语，查找修饰该词的正向情感动词或负向情感动词，情感得分为 *Score*。

②若满足条件①，则继续寻找并统计该专有名词词汇前面 5 个词语及该词汇后两个词语的范围内否定词的个数 $reverseTime_{SpeNoun}$，如果 $reverseTime_{SpeNoun}$ 是奇数，修正情感极性，令 $reverse_{SpeNoun}=-1$，否则不需要修正情感极性。

③寻找该专有名词前 5 个词语及该专有名词后两个词语范围内的程度级别词汇，用 *degree* 乘以匹配程度词语在程度级别词典中的情感强度系数，如公式(3)所示。

$$degree = degree \times DegDic_{DegWord} \tag{3}$$

④分别累加对应的正、负面情感得分，规则如公式(4)所示，当 $Score \times reverse_{SpeNoun} > 0$，情感得分累计到 *positive*，否则情感得分累计到 *negative*。

$$SentimentValue = \sum Score \times degree \times reverse_{SpeNoun} \tag{4}$$

(4) 文本谣言识别

通过文本情感值计算模块，对于每条文本，均有一个情感评分结果 *S* =(*positive*, *negative*)。文本长度、情感词语密度会影响 *S* 的值，需要归一化最初情感计算得分，如公式(5)–公式(7)所示。

假设最后每条文本的情感得分:

$$S = (S_{pos}, S_{neg}) \tag{5}$$

其中，

$$S_{pos}=\frac{|positive|}{|positive|+|negative|} \quad (6)$$

而 S_{neg} 可用 S_{pos} 表示:

$$S_{neg}=1-S_{pos} \quad (7)$$

因此 S_{neg} 和 S_{pos} 两者从不同方向表达情感程度, 本文仅以 S_{pos} 表示文本情感得分。对于不同文本 i 和 j, 规定它们之间的情感差异值 D 如公式(8)所示。

$$D=|S_{pos}^{i}-S_{pos}^{j}| \quad (8)$$

对于每个谣言案例来说, 设同等质量信息源(高质量或低质量)的信息源文本有 n 条, 那么对应的质量信息源情感得分为 S_{pos}^{Q}, 如公式(9)所示。

$$S_{pos}^{Q}=\frac{1}{n}\sum_{i=1}^{n}S_{pos}^{i} \quad (9)$$

其中, Q 取值为 H 代表高质量信息源文本, Q 取 L 代表低质量信息源文本。

需要说明的是, 存在一些文本总情感得分较小的情况, 表明文本的情感词命中太少, 情感倾向难以判断。本文设定: 如果情感评分总分$|positive|+|negative|<10$, 则该文本将不计入最后的倾向计算。

将每个谣言案例最终得到的高质量信息源得分与低质量信息源得分做情感差异评判, 对于高、低质量信息源得分差异 D', 如公式(10)所示。

$$D'=|S_{pos}^{H}-S_{pos}^{L}| \quad (10)$$

①当 $D'\leqslant\alpha$ 时, 认为高、低信息源在同一个谣言主题的核心事物上情感倾向一致, 认为低质量信息源的信息不属于谣言;

②当 $\alpha<D'<\beta$ 时, 认为两个文本之间的倾向大体一致, 低质量信息源可能存在一定的情感夸张, 但不属于谣言;

③当 $D'\geqslant\beta$ 时, 表示低质量信息源情感与高质量信息源发生冲突, 认为低质量信息源的说法属于谣言。

α、β 的具体取值由实验素材以及具体谣言主题决定, 目的是最大程度地划分情感的一致程度。

4　实验及结果分析评价

4.1　高低质量信息源界定

本文采用直接评价法界定信息源质量的高低, 评价指标包括文本错误率、编辑者身份、是否具有审核、发布、反馈和举报机制。

综合以上评价指标, 选定维基百科[19]、知乎[20]和果壳网[21]三个平台作为高质量信息源, 这三个网站的共同点为具有审核机制、编辑者为领域专业者、具有举报机制和评审机制等, 其质量较高, 其中维基百科在重大错误的数量上与《大英百科全书》几乎相等[22], 具有权威性。利用搜索引擎(如百度[23]、搜狗[24]等)搜索出的相关主题结果(除去来源为以上三个网站的结果), 作为低质量信息源。

4.2　高低质量信息来源

"流言百科[25]"是由果壳网开发的一个较为权威的辟谣平台, 在此辟谣平台上, 符合本文制定的"简单谣言"标准的谣言在"食品养生"、"医学健康"这两个领域的数量最多。从流言百科上寻找与"健康类"和"养生类"的文本内容相关的谣言案例, 关于该谣言的高质量文本, 则从果壳网、知乎、维基中通过关键词搜寻而来, 对于每个谣言案例, 要求以上三个网站至少有一处含有相关文本内容; 低质量文本则从普通搜索引擎搜到(剔除引用或直接来自以上三个网站的内容)的结果中获得, 对每个谣言案例, 要求至少有 4 条不重复文本。

在 2017 年 1 月至 2017 年 3 月采集数据, 由于低质量信息源各不相同, 暂未实现数据自动抓取, 因此采用人工收集的手段; 而高质量信息源关于特定对象的内容较少重复论述, 所以数据量偏少。在检索过程中, 大量谣言对象在高质量信息源没有搜索结果, 或者在低质量信息源处不满足数量要求, 舍弃, 导致抓取一个谣言案例需要耗费较大的检索资源与数据整理时间。最终收集了 30 个谣言案例, 共有 232 条数据参与实验, 其中有 48 条高质量信息、184 条低质量信息, 平均每个疑似谣言的案例有 1.6 条高质量信息, 6.1 条低质量信息。

4.3　情感词典建立

本文构建通用情感词典、专有名词词典、情感动词词典、否定词表和修饰词表以供情感词匹配计算情感值。基于大连理工大学中文情感词汇本体库[7]建立通用情感词典, 该情感词典共收录 27 386 个情感词汇, 并且具有词汇情感强度、情感类别和极性标注。该本体库的标注中, 0 代表中性, 1 代表褒义, 2 代表贬义, 3 代表兼有褒贬两性。每个词的情感强度分为 1, 3, 5, 7, 9 五档, 9 表示强度最大, 1 为强度最小。

本文将情感极性为 1 的 11 229 个词语纳入积极情感词典, 情感极性为 2 的 10 783 个词语纳入消极情感词典。其次, 对于情感极性为 0 的 5 374 个词语, 将其

中情感表示"乐"、"好"的词语纳入积极情感词典，将情感表示"怒"、"哀"、"惧"、"恶"、"惊"的词语纳入消极情感词典。而情感极性为 3 的词语，情感极性复杂，总计只有 78 个词汇，舍弃。每个词的情感评分取本体库中对应的情感强度值。

在实验过程中，人工补充了通用情感词中没有的词，如"明目"、"不法分子"等。评分标准采用对比评分法，如："奸商"与"不法分子"的情感倾向类似，"奸商"的评分在通用情感词典中为–5 分，而"不法分子"的批评意味比"奸商"更强烈，故予"不法分子"的评分是–7 分。

在建立领域专有名词词典方面，由于本文选取的信息来源文本多为养生类、疾病类和食品类与人体健康有关的疑似谣言文本。因此建立的词典为与人体疾病相关的专有名词词典，基于百度文库专业资料分类下的《疾病名称大全》[26]，通过文本处理过程中的观察与查缺补漏，人工补充了 52 个常见的疾病主题的词语。而相关正、负面情感动词词典则是对疾病有治疗作用或加剧作用的动词，如治愈表示治疗作用，而加重表示加剧作用。此外，本文结合已有情感分析研究及文本预处理过程的查漏补充，建立了本实验的程度副词词表和否定词表。

4.4 谣言识别结果

根据第 3 节的方法设计，利用 4.3 节建立的情感词典，计算同一谣言主题的案例下的不同质量来源谣言文本的情感差异，其中情感差异阈值 α 取 0.1、β 取 0.3。最终得到的谣言识别结果如表 1 所示。

表 1 谣言识别结果

	实际是谣言	实际不是谣言
预测是谣言	15	0
预测不是谣言	6	8

疑似谣言共计 30 条，实验共有 23 条谣言的结果判断正确，其中预测和人工判断都属于谣言的共计 15 条，预测和人工判断都不属于谣言的共计 8 条，6 条谣言的结果判断错误，1 条谣言的结果存在争议，排除在总结果之外，详细的判定结果可参看文末支撑数据清单。本文将选取典型判断正确文本及其实例、典型判断错误文本及其实例和存在争议谣言进行讨论。

(1) 典型判断正确情景及实例

典型判断正确的情景为：机器判断与人工判断结果一致，高质量信息源与低质量信息源对于同一件事物的情感倾向有明显的不同。例如低质量信息源的疑似谣言文本"牛奶有巨大危害"表示牛奶危害健康，人工判断情感为负面，机器判断情感得分为 0.354；而高质量信息源表示牛奶有助于预防骨质疏松，人工判断情感为正面，机器判断情感得分为 0.885。人工认为情感不一致，属于谣言；而机器计算情感差异 D' 值为 0.531>0.3，认为该条属于谣言。

(2) 典型判断错误情境及实例

一般而言，低质量信息源的情感直接，情感倾向明显；而高质量信息源则经常出现推理，转折的文本(如先提出可能存在的问题，然后在后续的文本中通过举例或者推理的方式给予解答)。现有的处理方法无法检测出段落之间的互相否定关系，导致高质量信息源的情感判断失误，引起错判。如："阿斯巴甜致癌，食用危害大"这条谣言，低质量信息源表示阿斯巴甜会损害神经系统，导致记忆力衰退、视力消失等症状；高质量信息源则表示"阿斯巴甜作为添加剂使用是安全的"。从以上观点来看，不论从机器通过情感词典匹配，还是人工判断，高质量信息源都属于正向情感。然而，高质量信息源中辟谣文本特征是：先提出存在"有许多指控声称阿斯巴甜的神经毒性作用，导致神经或精神症状"的谣言，再另起一段进行辟谣性的解释，导致负面情感虚高，机器判错。

(3) 存在争议谣言解释

以高质量信息源的信息作为标准，故而当不同的高质量信息源之间信息存在冲突时，可能是该条文本指示的知识在科学上暂时没有一个统一的说法，无法通过人工判定文本所述内容是否属于谣言。

本实验中，"蔓越莓能预防泌尿道疾病"这条疑似谣言的人工判断结果有争议，因为高质量信息源之间情感发生了冲突：果壳网摘录相关研究表明蔓越莓对泌尿道疾病可能有一定作用，但还没有具体证据支持，况且直接靠蔓越莓治病是不切实际的，属于负面情感；相反，维基百科则表示为"蔓越莓汁已被证实可有效降低心血管疾病、牙周病、胃溃疡与癌症等疾病的罹患风险"，属于正面情感。这说明，蔓越莓的治病功效可能在医学上尚有争议，并没有一个统一的判断标准。人工无法判断高质量信息源的总体情感倾向，故该文本为争议文本。

4.5　谣言识别结果评价

在结果评价中，本文参阅文献[27]所采用的查全率、查准率和 F 值用于评价此次实验结果，具体如表 2 和公式(11)–公式(16)所示。

表 2　谣言检测分类性能评价列表

	实际是谣言	实际不是谣言
预测是谣言	A	B
预测不是谣言	C	D

$$谣言文本查准率\ P_r = \frac{A}{A+B} \tag{11}$$

$$谣言文本查全率\ R_r = \frac{A}{A+C} \tag{12}$$

$$谣言文本\ F\ 值\ F1_r = \frac{2R_r \times P_r}{R_r + P_r} \tag{13}$$

$$非谣言文本查准率\ P_n = \frac{D}{C+D} \tag{14}$$

$$非谣言文本查全率\ R_n = \frac{D}{B+D} \tag{15}$$

$$非谣言文本\ F\ 值\ F1_n = \frac{2R_n \times P_n}{R_n + P_n} \tag{16}$$

由于存在争议的文本无法判断正确性，因此在统计中将其剔除，最终该实验的评价指标计算结果如表 3 所示。

表 3　谣言检测分类性能评价结果

指标	值
P_r	100%
R_r	71.42%
$F1_r$	83.34%
P_n	57.14%
R_n	100%
$F1_n$	72.73%

由表 3 可以看到，基于高、低信息源情感的一致性提出的识别谣言方法，在谣言文本预测上的 F 值是 83.34%，查全率为 71.42%，查准率为 100%；在非谣言文本预测上的 F 值为 72.73%，查全率为 100%，查准率为 57.14%。其中，在非谣言文本预测上查准率较低，其原因主要是非谣言文本具有严谨的推理逻辑，会对研究对象各个属性进行性质说明。本文提出在对疑似谣言文本的情感计算过程中，并没有提取某对象特定属性的情感值，由此导致相关判定失误。今后的研究可采用更加细粒度的情感分析技术进行谣言识别。

此实验评估结果表明，可用基于高、低质量信息源情感一致性计算识别特定主题的谣言。

5　结　语

基于情感强烈的谣言与非谣言之间存在情感冲突的理论依据，本文提出基于情感分析方法识别简单谣言的方法，该方法包括 4 个模块：高低质量信息源界定、文本预处理、文本情感值计算和文本谣言识别。为了验证方法的适用性，采用直接评价法界定高低质量信息源，界定知乎、维基百科和果壳网的为高质量信息源，人工抓取 30 个健康类、养生类领域谣言案例文本。进一步构建通用情感词典、专有名词词典等情感值计算的辅助词典，计算疑似谣言文本的情感值。通过计算来自高、低质量信息源的关于同一主题的疑似谣言文本的情感差异值，判定来自低质量信息源的文本是否属于谣言，最终正确判断了 23 个谣言案例。在方法评估上，本文提出的谣言识别方法能够较好地识别谣言。

本文存在的不足之处在于设定的情感计算方法在简单谣言识别上效果良好，但是对于语法规则复杂的疑似谣言文本会存在误判情况。今后的研究将考虑实现不同来源的信息源的文本自动抓取，增加复杂的情感计算规则，以识别更加复杂的谣言。

参考文献：

[1] 郭小安. 当代中国网络谣言的社会心理研究[M]. 北京: 中国社会科学出版社, 2015. (Guo Xiaoan. A Study on the Social Psychology of Contemporary Chinese Rumors Online[M]. Beijing: China Social Sciences Press, 2015.)

[2] 中国社会科学语言研究所词典编辑室. 现代汉语词典[M]. 第 5 版. 北京: 商务印书馆, 2005. (Dictionary of Chinese Social Sciences Language Research Institute.The Modern Chinese Dictionary [M]. The 5th Edition. Beijing: The Commercial Press, 2005.)

[3] 沙莲香. 社会心理学[M]. 第 2 版. 北京: 中国人民大学出版社, 2006. (Sha Lianxiang. Social Psychology[M]. The 2nd Edition. Beijing: China Renmin University Press, 2016.)

[4] Zhang L, Liu B. Sentiment Analysis and Opinion Mining[J]. Synthesis Lectures on Human Language Technologies, 2012, 30(1): 152-153.

[5] 杨立公, 朱俭, 汤世平.文本情感分析综述[J]. 计算机应用, 2013, 33(6): 1574-1578, 1607. (Yang Ligong, Zhu Jian, Tang Shiping. Survey of Text Sentiment Analysis[J]. Journal of Computer Applications, 2013, 33(6): 1574-1578, 1607.)

[6] Tong R M. An Operational System for Detecting and Tracking Opinions in Online Discussion [C]// Proceedings of the 26th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2001.

[7] 徐琳宏, 林鸿飞, 潘宇, 等.情感词汇本体的构造[J]. 情报学报, 2008, 27(2): 180-185. (Xu Linhong, Lin Hongfei, Pan Yu, et al. Constructing the Affective Lexicon Ontology[J]. Journal of the China Society for Scientific and Technical Information, 2008, 27(2): 180-185.)

[8] HowNet [EB/OL]. [2017-02-01]. http://www.keenage.com/html/e_index.html.

[9] 陈晓东. 基于情感词典的中文微博情感倾向分析研究[D]. 武汉: 华中科技大学, 2012. (Chen Xiaodong. Research on Sentiment Dictionary Based Emotional Tendency Analysis of Chinese MicroBlog[D]. Wuhan: Huazhong University of Science & Technology, 2012.)

[10] 肖璐, 陈果, 刘继云. 基于情感分析的企业产品级竞争对手识别研究——以用户评论为数据源[J]. 图书情报工作, 2016, 60(1): 83-90, 97. (Xiao Lu, Chen Guo, Liu Jiyun. Study on Identification of Enterprise Product Level Competitor Based on Sentiment Analysis: Taking User Reviews for Data Resources[J]. Library and Information Service, 2016, 60(1): 83-90, 97.)

[11] Litton I. # TwitterCritic: Sentiment Analysis of Tweets to Predict TV Ratings [EB/OL]. [2017-03-05]. http://digitalcommons.calpoly. edu/cgi/viewcontent.cgi?article=1057&context=statsp.

[12] Nguyen T H, Shirai K. Topic Modeling Based Sentiment Analysis on Social Media for Stock Market Prediction[C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. 2015: 1354-1364.

[13] 毛二松, 陈刚, 刘欣, 等. 基于深层特征和集成分类器的微博谣言检测研究[J]. 计算机应用研究, 2016, 33(11): 3369-3373. (Mao Ersong, Chen Gang, Liu Xin, et al. Research on Detecting Micro-blog Rumors Based on Deep Features and Ensemble Classifier[J]. Application Research of Computers, 2016, 33(11): 3369-3373.)

[14] Qazvinian V, Rosengren E, Radev D R, et al. Rumor Has It: Identifying Misinformation in Microblogs[C]// Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. 2012: 1589-1599.

[15] Kwon S, Cha M, Jung K. Rumor Detection over Varying Time Windows[J]. PLoS One, 2017, 12(1): e0168344.

[16] 张志安, 束开荣, 何凌南. 微信谣言的主题与特征[J]. 新闻与写作, 2016(1): 60-64. (Zhang Zhian, Shu Kairong, He Lingnan. The Topics and Features About Rumors on the WeChat[J]. News and Writing, 2016(1): 60-64.)

[17] 马费成, 宋恩梅. 信息管理学基础[M]. 第 2 版. 武汉: 武汉大学出版社, 2011: 136-142. (Ma Feicheng, Song Enmei. Principles of Information Management [M]. The 2nd Edition. Wuhan: Wuhan University Press, 2011: 136-142.)

[18] 杜嘉忠, 徐健, 刘颖. 网络商品评论的特征–情感词本体构建与情感分析方法研究[J]. 现代图书情报技术, 2014(5): 74-82. (Du Jiazhong, Xu Jian, Liu Ying. Research on Construction of Feature-Sentiment Ontology and Sentiment Analysis[J]. New Technology of Library and Information Service, 2014(5): 74-82.)

[19] 维基百科 [EB/OL]. [2017-03-26]. https://www.wikipedia.org/. (Wikipedia[EB/OL].[2017-03-26]. https://www.wikipedia.org/.)

[20] 知乎[EB/OL]. [2017-03-26]. https://www.zhihu.com. (Zhihu [EB/OL]. [2017-03-26]. https://www.zhihu.com.)

[21] 果壳网 [EB/OL]. [2017-03-26]. http://www.guokr.com/. (Guokr [EB/OL]. [2017-03-26]. http://www.guokr.com/.)

[22] Giles J. Internet Encyclopedias Go Head to Head[J]. Nature, 2005, 138(15): 900-901.

[23] 百度[EB/OL]. [2017-03-26]. https://www.baidu.com. (Baidu [EB/OL]. [2017-03-26]. https: //www.baidu.com.)

[24] 搜狗 [EB/OL]. [2017-03-26]. https://www.sogou.com/. (Sougou [EB/OL]. [2017-03-26]. https://www.sogou.com/.)

[25] 流言百科 [EB/OL]. [2017-03-26]. http://www.liuyanbaike.com/. (Liuyanbaike [EB/OL]. [2017-03-26]. http://www.liuyanbaike.com/.)

[26] 百度文库.疾病名称大全 [EB/OL]. [2017-03-26]. https://wenku.baidu.com/view/0cd82214a5e9856a5712601f.Html. (Baidu Wenku. Full Listing of Disease Name [EB/OL]. [2017-03-26]. https://wenku.baidu.com/view/0cd82214a5e9856a5712601f.html.)

[27] 吴江, 唐常杰, 李太勇, 等. 基于语义规则的 Web 金融文本情感分析[J]. 计算机应用, 2014, 34(2): 481-485. (Wu Jiang, Tang Changjie, Li Taiyong, et al. Sentiment Analysis on Web Financial Text Based on Semantic Rules[J]. Journal of Computer Applications, 2014, 34(2): 481-485.)

作者贡献声明:

首欢容: 算法设计, 采集数据, 完成实验, 撰写论文初稿;
邓淑卿: 提出部分修改意见, 数据整理, 论文修改;
徐健: 提出研究思路, 设计研究方案, 论文最终版本修订。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail: 674001239@qq.com。

[1] 首欢容, 邓淑卿, 徐健. 数据.zip. 实验中所有谣言案例的数据集.
[2] 首欢容, 邓淑卿, 徐健. 词典.zip. 本实验建立的所有词典.
[3] 首欢容, 邓淑卿, 徐健. 实验结果详情汇总表格.xlsx. 实验结果详情汇总.

收稿日期: 2017-05-27
收修改稿日期: 2017-07-15

Detecting Online Rumors with Sentiment Analysis

Shou Huanrong Deng Shuqing Xu Jian
(School of Information Management, Sun Yat-Sen University, Guangzhou 510006, China)

Abstract: [**Objective**] This paper aims to identify rumors automatically with the help of sentiment analysis. [**Methods**] First, we chose high-quality and low-quality information sources. Then, we calculated the sentiment value and difference between the information from different sources. Based on the assumption that the information from high-quality source was more reliable, information from low-quality channels could be listed as rumor if the sentiment difference between them exceeded the pre-set threshold. [**Results**] We applied the proposed method to information on food and health as well as health and medical issues, and then successfully identified twenty-three rumors from thirty suspected cases. The accuracy rate of rumor detection was 76.67%, the F-value was 83.34%, the recall and precision was 71.42% and 100%, respectively. For non-rumor message, the F-value, recall, and precision were 72.73%, 100% and 57.14%. [**Limitations**] We did not extract the data automatically from different sources and the sample size was relatively small. [**Conclusions**] Sentiment analysis could help us identify rumors effectively.

Keywords: Sentiment Analysis Sentiment Lexicon Rumor Identification Rumor Detection